EXHIBIT B

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, A01K 67/027, C12N 15/12

A1

(11) International Publication Number: WO 99/29896

(43) International Publication Date: 17 June 1999 (17.06.99)

(21) International Application Number: PCT/CA98/01133

(22) International Filing Date: 7 December 1998 (07.12.98)

(30) Priority Data:
2,218,199 9 December 1997 (09.12.97) CA

(71) Applicant *(for all designated States except US)*: McGILL UNIVERSITY [CA/CA]; 845 Sherbrooke Street West, Montréal, Québec H3A 2T5 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ROULEAU, Guy, A. [CA/CA]; Appartement 7, 4850 Côte Saint–Luc, Montréal, Québec H3W 2H2 (CA). BRAIS, Bernard [CA/CA]; 745 Champagneur, Outremont, Québec H2V 3P9 (CA).

(74) Agents: COTE, France et al.; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: SHORT GCG EXPANSIONS IN THE *PAB II* GENE FOR OCULOPHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC THEREOF

(57) Abstract

The present invention relates to a human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in SEQ ID NO:3, which includes introns and flanking genomic sequence. The allelic variants of GCG repeat of the human PAB II gene are associated with a disease related with protein accumulation in nucleus, such as polyalanine accumulation, a disease related with swallowing difficulties, such as oculopharyngeal muscular dystrophy. The present invention also relates to a method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of: a) obtaining a nucleic acid sample of said patient; and b) determining allelic variants of GCG repeat of the gene of claim 1, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SHORT GCG EXPANSIONS IN THE *PAB II* GENE FOR OCULO-PHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC THEREOF

BACKGROUND OF THE INVENTION

(a) Field of the Invention

The invention relates to PAB II gene, and its uses thereof for the diagnosis, prognosis and treatment of a disease related with protein accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

(b) Description of Prior Art

Autosomal dominant oculopharyngeal muscular dystrophy(OPMD) is an adult-onset disease with a world-wide distribution. It usually presents in the sixth decade with progressive swallowing difficulties (dysphagia), eye lid drooping (ptosis) and proximal limb weakness. Unique nuclear filament inclusions in skeletal muscle fibers are its pathological hallmark (Tome, F.M.S. & Fardeau, *Acta Neuropath.* **49**, 85-87 (1980)). We isolated the poly(A) binding protein II (PAB II) gene from a 217 kb candidate interval in chromosome 14q11. A (GCG)6 repeat encoding a polyalanine tract located at the N-terminus of the protein was expanded to (GCG)8-13 in the 144 OPMD families screened. More severe phenotypes were observed in compound heterozygotes for the (GCG)9 mutation and a (GCG)7 allele found in 2% of the population, whereas homozygosity for the (GCG)7 allele leads to autosomal recessive OPMD. Thus the (GCG)7 allele is an example of a polymorphism which can act as either a modifier of a dominant phenotype or as a recessive mutation. Pathological expansions of the polyalanine tract may cause mutated PAB II oligomers to accumulate as filament inclusions in nuclei.

It would be highly desirable to be provided with a tool for the diagnosis, prognosis and treatment of a disease related with polyalanine accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

SUMMARY OF THE INVENTION

One aim of the present invention is to provide a tool for the diagnosis, prognosis and treatment of a disease related with polyalanine accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

In accordance with the present invention there is provided a human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in Fig. 4, which includes introns and flanking genomic sequence.

The allelic variants of GCG repeat of the human PAB II gene are associated with a disease related with protein accumulation in nucleus, such as polyalanine accumulation, or with a disease related with swallowing difficulties, such as oculopharyngeal muscular dystrophy.

In accordance with the present invention there is also provided a method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of:

a) obtaining a nucleic acid sample of said patient; and

b) determining allelic variants of GCG repeat of the gene of the human PAB II gene, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus, such as polyalanine accumulation and oculopharyngeal muscular dystrophy.

The long allelic variants have from about 245 to about 263 bp in length.

In accordance with the present invention there is also provided a non-human mammal model for the PAB II gene of the human PAB II gene, whose germ cells and somatic cells are modified to express at least one allelic variant of the PAB II gene and wherein said

allelic variant of the PAB II being introduced into the mammal, or an ancestor of the mammal, at an embryonic stage.

In accordance with the present invention there is also provided a method for the screening of therapeutic agents for the prevention and/or treatment of oculopharyngeal muscular dystrophy, which comprises the steps of:

a) administering said therapeutic agents to the non-human mammal of the present invention or oculopharyngeal muscular dystrophy patients; and

b) evaluating the prevention and/or treatment of development of oculopharyngeal muscular dystrophy in said mammal or said patients.

In accordance with the present invention there is also provided a method to identify genes part of or interacting with a biochemical pathway affected by PAB II gene, which comprises the steps of:

a) designing probes and/or primers using the hGT1 gene of the PAB II gene and screening oculopharyngeal muscular dystrophy patients samples with said probes and/or primers; and

b) evaluating the identified gene role in oculopharyngeal muscular dystrophy patients.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A-B illustrate the positional cloning of the *PAB II* gene;

Figs. 2A-G illustrate the OPMD $(GCG)_n$ expansion sizes and sequence of mutations (SEQ ID NOS:1-2);

Fig. 3 illustrates the age distribution of swallowing time (*st*) for French Canadian OPMD carriers of the $(GCG)_9$ mutation; and

Fig. 4 illustrates the nucleotide sequence of human poly(A) binding protein II (hPAB II)(SEQ ID NO:3).

DETAILED DESCRIPTION OF THE INVENTION

In order to identify the gene mutated in OPMD, we constructed a 350 kb cosmid contig between flanking markers D14S990 and D14S1457 (Fig. 1A). Positions of the *PAB II* selected cDNA clones in relation to the *Eco*RI restriction map and the Genealogy-based Estimate of Historical Meiosis (GEHM)-derived candidate interval (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)).

The human poly(A) binding protein II gene (PAB II) is encoded by the nucleotide sequence as set forth in Fig. 4.

Twenty-five cDNAs were isolated by cDNA selection from the candidate interval (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)). Three of these hybridized to a common 20 kb EcoRI restriction fragment and showed high sequence homology to the bovine poly(A) binding protein II gene(bPAB II) (Fig. 1A). The PAB II gene appeared to be a good candidate for OPMD because it mapped to the genetically defined 0.26 cM candidate interval in 14q11 (Fig. 1A), its mRNA showed a high level of expression in skeletal muscle, and the PAB II protein is exclusively localized to the nucleus (Krause, S. et al., *Exp. Cell Res.* **214**, 75-82 (1994)) where it acts as a factor in mRNA polyadenylation (Whale, E., *Cell* **66**, 759-768 (1991); Whale, E. et al., *J. Biol. Chem.* **268**,

2937-2945 (1993); Bienroth, S. et al., EMBO J. 12, 585-594 (1993)).

We subcloned a 8 kb HindIII genomic fragment containing the PAB II gene, and sequenced 6002 bp (GenBank: AF026029)(Nemeth, A. et al., *Nucleic Acids Res.* 23, 4034-4041 (1995)) (Fig. 1B). Genomic structure of the *PAB II* gene, and position of the OPMD $(GCG)_n$ expansions. Exons are numbered. Introns 1 and 6 are variably present in 60% of cDNA clones. ORF, open reading frame; cen, centromere and tel, telomere.

The coding sequence was based on the previously published bovine sequence (GenBank: X89969) and the sequence of 31 human cDNAs and ESTs. The gene is composed of 7 exons and is transcribed in the cen-qter orientation (Fig. 1B). Multiple splice variants are found in ESTs and on Northern blots (Nemeth, A. et al., *Nucleic Acids Res.* 23, 4034-4041 (1995)). In particular, introns 1 and 6 are present in more than 60% of clones (Fig. 1B)(Nemeth, A. et al., *Nucleic Acids Res.* 23, 4034-4041 (1995)). The coding and protein sequences are highly conserved between human, bovine and mouse (GenBank: U93050). 93% of the PAB II sequence was readily amenable to RT-PCR- or genomic-SSCP screening. No mutations were uncovered using both techniques. However, a 400 bp region of exon 1 containing the start codon could not be readily amplified. This region is 80% GC rich. It includes a (GCG)6 repeat which codes for the first six alanines of a homopolymeric stretch of 10 (Fig. 2G). Nucleotide sequence of the mutated region of *PAB II*. Amino acid sequences of the N-terminus polyalanine stretch and position of the OPMD alanine insertions.

Special conditions were designed to amplify by PCR a 242 bp genomic fragment including this GCG-repeat. The (GCG)6 allele was found in 98% of French

Canadian non-OPMD control chromosomes, whereas 2% of chromosomes carried a (GCG)7 polymorphism (n=86) (Brais, B. et al., *Hum. Mol. Genet.* 4, 429-434 (1995)).

Screening OPMD cases belonging to 144 families showed in all cases a PCR product larger by 6 to 21 bp than that found in controls (Fig. 2A). $(GCG)_6$ normal allele (N) and the six different $(GCG)_n$ expansions observed in 144 families.

Sequencing of these fragments revealed that the increased sizes were due to expansions of the GCG repeat (Fig. 2G). Fig. 2F shows the sequence of the (GCG)9 French Canadian expansion in a heterozygous parent and his homozygous child. Partial sequence of exon 1 in a normal $(GCG)_6$ control (N), a heterozygote (ht.) and a homozygote (hm.) for the $(GCG)_9$-repeat mutation. The number of families sharing the different (GCG)n-repeats expansions is shown in Table 1.

Table 1

Number of families sharing the different dominant $(GCG)_n$ OPMD mutations

Mutations	Polyalanine	Families
$(GCG)_8$	12	4
$(GCG)_9$	13	99
$(GCG)_{10}$	14	19
$(GCG)_{11}$	15	16
$(GCG)_{12}$	16	5
$(GCG)_{13}$	17	1
Total		**144**

†, 10 alanine residues in normal PAB II.

The (GCG)9 expansion shared by 70 French Canadian families is the most frequent mutation we observed (Table 1). The (GCG)9 expansion is quite stable, with a single doubling observed in family F151 in an estimated 598 French Canadian meioses (Fig. 2C). The doubling of

the French Canadian $(GCG)_9$ expansion is demonstrated in Family F151.

This contrasts with the unstable nature of previously described disease-causing triplet-repeats (Rosenberg, R.N., *New Eng. J. Med.* **335**, 1222-1224 (1996)).

Genotyping of all the participants in our clinical study of French Canadian OPMD provided molecular insights into the clinical variability observed in this condition. The genotypes for both copies of the PAB II mutated region were added to an anonymous version of our clinical database of 176 $(GCG)_9$ mutation carriers (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995)). Severity of the phenotype can be assessed by the swallowing time (st) in seconds taken to drink 80 cc of ice-cold water (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995); Bouchard, J.-P. et al., *Can. J. Neurol. Sci.* **19**, 296-297 (1992)). The late onset and progressive nature of the muscular dystrophy is clearly illustrated in heterozygous carriers of the $(GCG)_9$ mutation (bold curve in Fig. 3) when compared the average st of control $(GCG)_6$ homozygous participants(n=76, thinner line in Fig. 3). The bold curve represents the average OPMD *st* for carriers of only one copy of the $(GCG)_9$ mutation (n=169), while the thinner line corresponds to the average *st* for $(GCG)_6$ homozygous normal controls(n=76). The black dot corresponds to the *st* value for individual VIII. Roman numerals refer to individual cases shown in Figs. *2B*, *2D* and discussed in the text. Genotype of a homozygous $(GCG)_9$ case and her parents (Fig. *2B*). Independent segregation of the $(GCG)_7$ allele. Case V has a more severe OPMD phenotype (Fig. *2D*).

Two groups of genotypically distinct OPMD cases have more severe swallowing difficulties. Individuals

I, II, and III have an early-onset disease and are homozygous for the (GCG)9 expansion ($P < 10^{-5}$) (Figs. 2B, F). Cases IV, V, VI and VII have more severe phenotypes and are compound heterozygotes for the (GCG)9 mutation and the (GCG)7 polymorphism ($P < 10^{-5}$). In Fig. 2D the independent segregation of the two alleles is shown. Case V, who inherited the French Canadian (GCG)9 mutation and the (GCG)7 polymorphism, is more symptomatic than his brother VIII who carries the (GCG)9 mutation and a normal (GCG)6 allele (Figs. 2D and 3). The (GCG)7 polymorphism thus appears to be a modifier of severity of dominant OPMD. Furthermore, the (GCG)7 allele can act as a recessive mutation. This was documented in the French patient IX who inherited two copies of the (GCG)7 polymorphism and has a late-onset autosomal recessive form of OPMD (Fig. 2E). Case IX, who has a recessive form of OPMD, is shown to have inherited two copies of the (GCG)7 polymorphism.

This is the first description of short trinucleotide repeat expansions causing a human disease. The addition of only two GCG repeats is sufficient to cause dominant OPMD. OPMD expansions do not share the cardinal features of "dynamic mutations". The GCG expansions are not only short they are also meiotically quite stable. Furthermore, there is a clear cut-off between the normal and abnormal alleles, a single GCG expansion causing a recessive phenotype. The PAB II (GCG)7 allele is the first example of a relatively frequent allele which can act as either a modifier of a dominant phenotype or as a recessive mutation. This dosage effect is reminiscent of the one observed in a homozygote for two dominant synpolydactyly mutations. In this case, the patient had more severe deformities because she inherited two duplications causing an expansion in the

polyalanine tract of the HOXD13 protein (Akarsu, A.N. et al., *Hum. Mol. Genet.* **5**, 945-952 (1996)). A duplication causing a similar polyalanine expansion in the a subunit 1 gene of the core-binding transcription factor (CBFα1) has also been found to cause dominant cleidocranial dysplasia (Mundlos, S. et al., *Cell* **89**, 773-779 (1997)). The mutations in these two rare diseases are not triplet-repeats. The are duplications of "cryptic repeats" composed of mixed synonymous codons and are thought to result from unequal crossing over (Warren, S.T., *Science* **275**, 408-409 (1997)). In the case of OPMD, slippage during replication causing a reiteration of the GCG codon is a more likely mechanism (Wells, D.R., *J. Biol. Chem.* **271**, 2875-2878 (1996)).

Different observations converge to suggest that a gain of function of PAB II may cause the accumulation of nuclear filaments observed in OPMD (Tome, F.M.S. & Fardeau, *Acta Neuropath.* **49**, 85-87 (1980)). PAB II is found mostly in dimeric and oligomeric form (Nemeth, A. et al., *Nucleic Acids Res.* **23**, 4034-4041 (1995)). It is possible that the polyalanine tract plays a role in polymerization. Polyalanine stretches have been found in many other nuclear proteins such as the HOX proteins, but their functions is still unknown (Davies, S.W. et al., *Cell* **90**, 537-548 (1997)). Alanine is a highly hydrophobic amino acid present in the cores of proteins. In dragline spider silk, polyalanine stretches are thought to form B-sheet structures important in ensuring the fibers' strength (Simmons, A.H. et al., *Science* **271**, 84-87 (1996)). Polyalanine oligomers have also been shown to be extremely resistant to chemical denaturation and enzymatic degradation (Forood, B. et al., *Bioch. and Biophy. Res. Com.* **211**, 7-13 (1995)). One can speculate that PAB II oligomers comprised of a sufficient number of mutated molecules

might accumulate in the nuclei by forming undegradable polyalanine rich macromolecules. The rate of the accumulation would then depend on the ratio of mutated to non-mutated protein. The more severe phenotypes observed in homozygotes for the (GCG)9 mutations and compound heterozygotes for the (GCG)9 mutation and (GCG)7 allele may correspond to the fact that in these cases PAB II oligomers are composed only of mutated proteins. The ensuing faster filament accumulation could cause accelerated cell death. The recent description of nuclear filament inclusions in Huntington's disease, raises the possibility that "nuclear toxicity" caused by the accumulation of mutated homopolymeric domains is involved in the molecular pathophysiology of other triplet-repeat diseases (Davies, S.W. et al., *Cell* **90**, 537-548 (1997); Scherzinger, E. et al., *Cell* **90**, 549-558 (1997); DiFiglia, M. et al., *Science* **277**, 1990-1993 (1997)). Future immunocytochemical and expression studies will be able to test this pathophysiological hypothesis and provide some insight into why certain muscle groups are more affected while all tissues express PAB II.

Methods

Contig and cDNA selection

The cosmid contig was constructed by standard cosmid walking techniques using a gridded chromosome 14-specific cosmid library (Evans, G.A. et al., *Gene* **79**, 9-20 (1989)). The cDNA clones were isolated by cDNA selection as previously described (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)).

Cloning of the PAB II gene. Three cDNA clones corresponding to PAB II were sequenced (Sequenase,

USB). Clones were verified to map to cosmids by Southern hybridization. The 8 kb HindIII restriction fragment was subcloned from cosmid 166G8 into pBluescriptII (SK) (Stratagene). The clone was sequenced using primers derived from the bPABII gene and human EST sequences. Sequencing of the PAB II introns was done by primer walking.

PAB II mutation screening and sequencing. All cases were diagnosed as having OPMD on clinical grounds (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995)). RT-PCR- and genomic SSCP analyses were done using standard protocols (Lafrenière, R.G. et al., *Nat. Genet.* **15**, 298-302 (1997)). The primers used to amplify the PAB II mutated region were: 5'-CGCAGTGCCCCGCCTTAGA-3' (SEQ ID NO:4) and 5'-ACAAGATGGCGCCGCCGCCCCGGC-3' (SEQ ID NO:5). PCR reactions were performed in a total volume of 15 ml containing: 40 ng of genomic DNA; 1.5 mg of BSA; 1 mM of each primer; 250 mM dCTP and dTTP; 25 mM dATP; 125 mM of dGTP and 125 mM of 7-deaza-dGTP (Pharmacia); 7.5% DMSO; 3.75 mCi[35S]dATP, 1.5 unit of Taq DNA polymerase and 1.5 mM $MgCl_2$ (Perkin Elmer). For non-radioactive PCR reactions the [35S]dATP was replaced by 225 mM of dATP. The amplification procedure consisted of an initial denaturation step at 95°C for five minutes, followed by 35 cycles of denaturation at 95°C for 15 s, annealing at 70°C for 30 s, elongation at 74°C for 30 s and a final elongation at 74°C for 7 min. Samples were loaded on 5% polyacrylamide denaturing gels. Following electrophoresis, gels were dried and autoradiographs were obtained. Sizes of the inserts were determined by comparing to a standard M13 sequence (Sequenase, USB). Fragments used for sequencing were gel-purified. Sequencing of the mutated fragment using the Amplicycle kit (Perkin Elmer) was done with the 5'-

CGCAGTGCCCCGCCTTAGAGGTG-3' (SEQ ID NO:6) primer at an elongation temperature of 68°C.

Stability of (GCG)-repeat expansions. The meiotic stability of the (GCG)9-repeat was estimated based on our large French Canadian OPMD cohort. We previously established that a single ancestral OPMD carrier chromosome was introduced in the French Canadian population by three sisters in 1648. Seventy of the seventy one French Canadian OPMD families tested to date segregate a (GCG)9 expansion. However, in family F151, the affected brother and sister, despite sharing the French Canadian ancestral haplotype, carry a (GCG)12 expansion twice the size of the ancestral (GCG)9 mutation (Fig. 2C). In our founder effect study, we estimated that 450 (304-594) historical meioses shaped the 123 OPMD cases belonging to 42 of the 71 enrolled families. Our screening of our full set of participants allowed us to identify another 148 (GCG)9 carrier chromosomes. Therefore, we estimate that a single mutation of the (GCG)9 expansion has occurred in 598 (452-742) meioses.

Genotype-phenotype correlations. 176 carriers of at least one copy of the (GCG)9 mutation were examined during the early stage of the linkage study. All were asked to swallow 80 cc of ice-cold water as rapidly as possible. Testing was stopped after 60 seconds. The swallowing time (st) was validated as a sensitive test to identify OPMD cases (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995); Bouchard, J.-P. et al., *Can. J. Neurol. Sci.* **19**, 296-297 (1992)). The st values for 76 (GCG)6 homozygotes normal controls is illustrated in Fig. 3. Analyses of variance were computed by two-way ANOVA (SYSTAT package). For the (GCG)9 homozygotes their mean st value was compared to the mean value for all (GCG)9 heterozygotes aged 35-40 ($P < 10^{-5}$). For the (GCG)9 and (GCG)7 compound heterozygotes their mean st

value was compared to the mean value for all (GCG)9 heterozygotes aged 45-65 ($P < 10^{-5}$).

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth, and as follows in the scope of the appended claims.

WHAT IS CLAIMED IS:

1. A human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in SEQ ID NO:3, which includes introns and flanking genomic sequence.

2. The gene of claim 1, wherein allelic variants of GCG repeat are associated with a disease related with protein accumulation in nucleus.

3. The gene of claim 2, wherein said protein accumulation is polyalanine accumulation.

4. The gene of claim 1, wherein allelic variants of GCG repeat are associated with a disease related with swallowing difficulties.

5. The gene of claim 1, wherein said disease is oculopharyngeal muscular dystrophy.

6. A method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of:

a) obtaining a nucleic acid sample of said patient; and

b) determining allelic variants of GCG repeat of the gene of claim 1, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus.

7. The method of claim 6, wherein said disease is oculopharyngeal muscular dystrophy.

8. The method of claim 7, wherein said long allelic variants have from about 245 to about 263 bp in length.

9. A non-human mammal model for the PAB II gene of claim 1, whose germ cells and somatic cells are modified to express at least one allelic variant of the PAB II gene and wherein said allelic variant of the PAB II being introduced into the mammal, or an ancestor of the mammal, at an embryonic stage.

10. A method for the screening of therapeutic agents for the prevention and/or treatment of oculopharyngeal muscular dystrophy, which comprises the steps of:

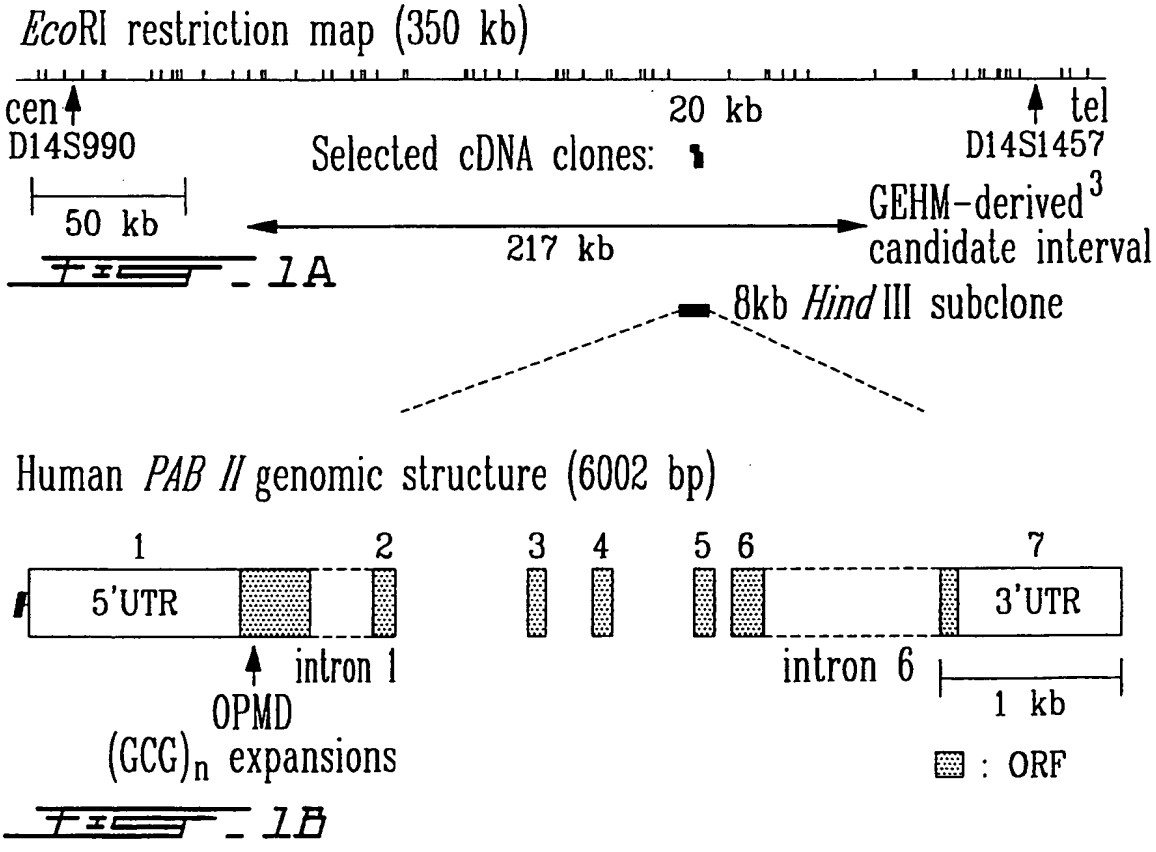
a) administering said therapeutic agents to the non-human mammal of claim 9 or oculopharyngeal muscular dystrophy patients; and

b) evaluating the prevention and/or treatment of development of oculopharyngeal muscular dystrophy in said mammal or said patients.

11. A method to identify genes part of or interacting with a biochemical pathway affected by PAB II gene, which comprises the steps of:

a) designing probes and/or primers using the hGT1 gene of claim 1 and screening oculopharyngeal muscular dystrophy patients samples with said probes and/or primers; and

b) evaluating the identified gene role in oculopharyngeal muscular dystrophy patients.


EcoRI restriction map (350 kb)
cen
D14S990
20 kb
tel
D14S1457
Selected cDNA clones:
50 kb
217 kb
GEHM-derived 3
candidate interval
1A
8kb HindIII subclone
Human PAB II genomic structure (6002 bp)
1
2
3
4
5
6
7
5'UTR
3'UTR
intron 1
intron 6
1 kb
OPMD
(GCG)n expansions
: ORF
1B

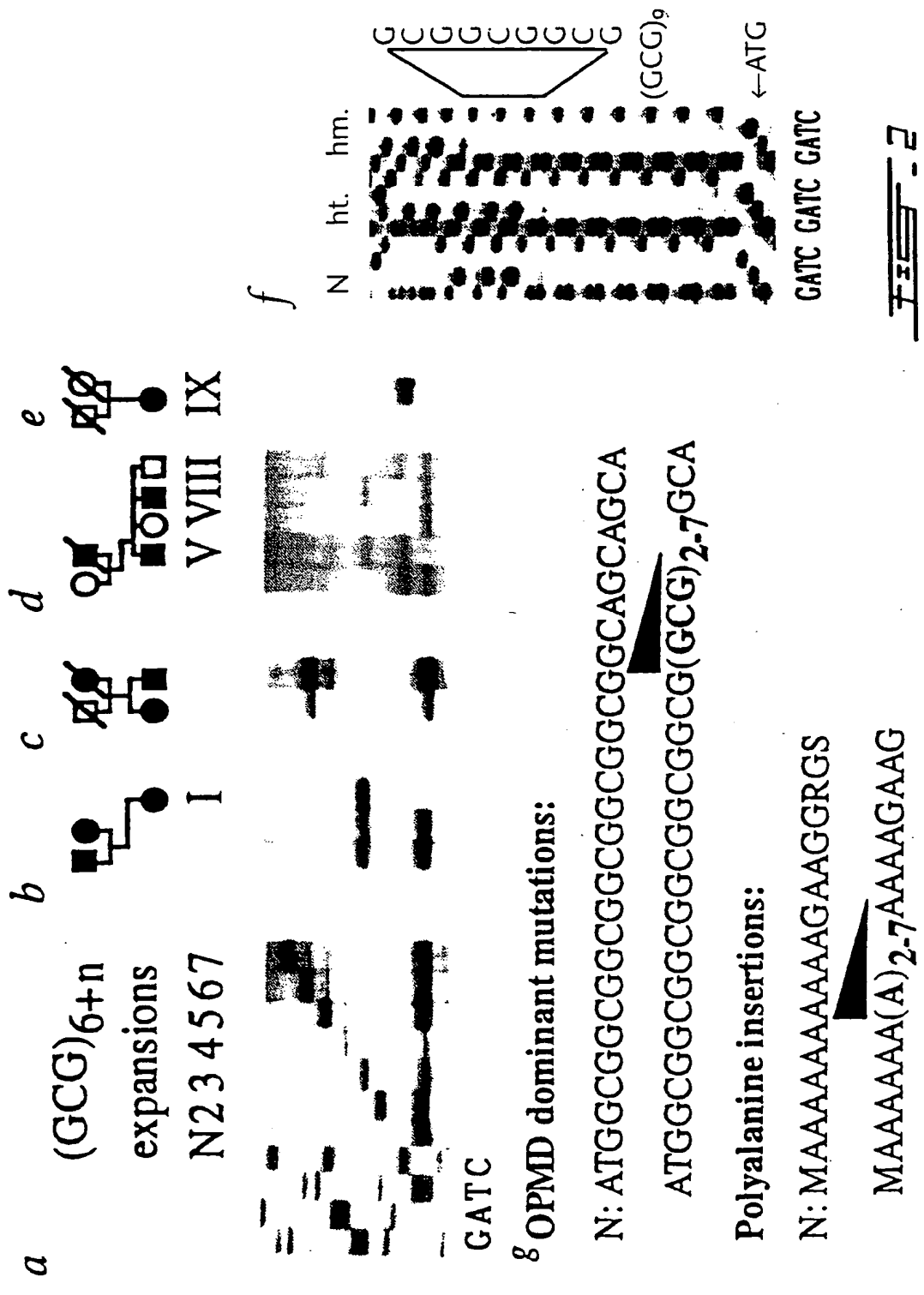

a
b
c
d
e
(GCG)6+n
expansions
N 2 3 4 5 6 7
I
V VIII IX
GATC
g OPMD dominant mutations:
N: ATGGCGGCGGCGGCGGCGGCAGCAGCA
ATGGCGGCGGCGGCGGCGGCG(GCG)2-7GCA
Polyalanine insertions:
N: MAAAAAAAAAAGAAGGRGS
MAAAAAA(A)2-7AAAAGAAG
f
N ht. hm.
(GCG)9
←ATG
GATC GATC GATC

Fig - 2

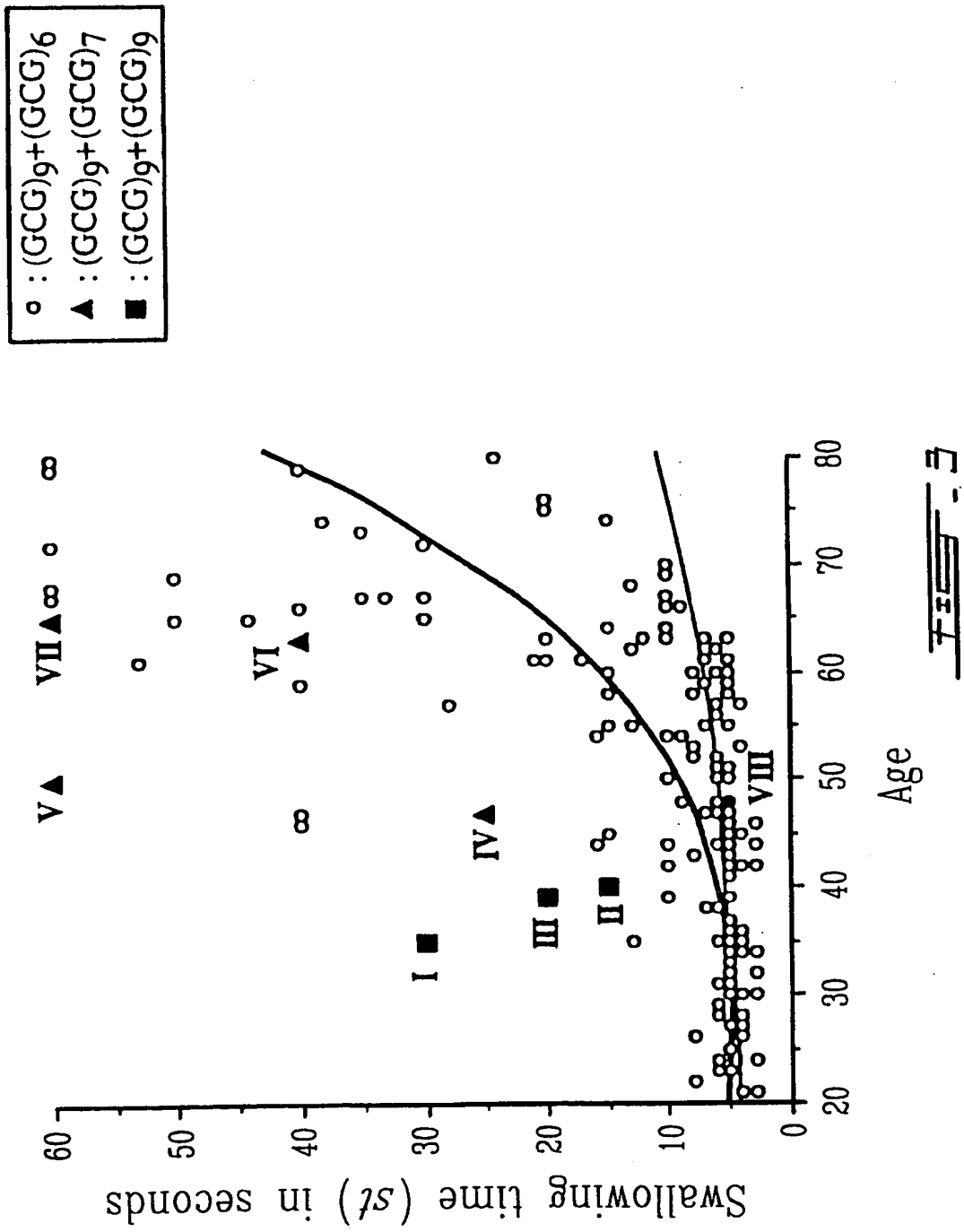

o : (GCG)9+(GCG)6
▲ : (GCG)9+(GCG)7
■ : (GCG)9+(GCG)9
Swallowing time (st) in seconds
Age
I
II
III
IV
V
VI
VII
VIII

Fig - 3

```
   1 aatgaaggtg gacacccaaa tagccccaat acaaatgcct gttcaatcaa ccaaacatct
  61 aagcagcaca tctatgtggt agcatattgc caggccgtga gactgcgaat ataaatagga
 121 accgccccctc atctgcaggc gctcacaacc tagttagcaa acagtaaaac aattaagcgc
 181 gccgtggaca taggcccact tgtcctgggа aatgagggga agctgggggtt tgcagtggtt
 241 tgattgaagg gggactacat gttagaggca cagactgggt gcaggtacac ccaaaggaac
 301 gagaagagtg gaaggaaaca acatccacaa agtaaccaca tgctggcgta tcgaaggccg
 361 tgatttacgg tttgagact ttacctcgcc agcaaagggg ggccagtctg ttagcggtgc
 421 agattggagg ggtgacattg gaagctgtcc aggaaaaaga aaatggaact ggggagcaga
 481 aggcctacgc aagagggcgg gacagacagg acttgtgact agtagctctg gactgaggaa
 541 tcctcccctgc ttctgggtgc gggagagcta gtggatgatg gtgccaataa cctggatggg
 601 gaaagtaagc tccctccctgg aatgcttcat tcacaacctc catttttcagc aacatcccat
 661 ctactggtgc ttcctgggtcg agatacaagt ttcctgaaac tgctgctctg ttttgggcct
 721 cacccggcca acagctcact agctggcaag cagtagtatc aagatggcgg ccccctagga
 781 ctggctagtc atgtgacctc gggtttccca agtttgaagc ccggcagtcc ttcgggggc
 841 aaggttcacc tgtcacgaaa cgagtgtcac ccttcgact ctcgcaagcc aatcggcatc
 901 tgagactggg ccactgcggt gaggcgatcg gaagattggt cctttccagt cgcctagcta
 961 gggccaatca cggagcgtcc catacttcgc gggcccgccc gtaggccggg gagaagcagg
1021 aatatcgtca cagcgtggcg gtattattac ctaaggactc gataggaggt gggacgcgtg
1081 ttgattgaca ggcagatttc cctaccgggа tttgagaatt tggcgcagtg cccgccttag
1141 aggtgcgctt atttgattgc caagtaatat tccccaatgg agtactagct catggtgacg
```

FIG - 4A

```
1201 ggcaggcagc ttgagctaat gagtcctccg tggccggcgc agctctccac atgccggggcg
1261 gcgggccccca gtctgagcgg cgatggcggc ggcggcgccg gcggcagcag cagcgggggc
1321 tgcgggcggt cgggctccg ggccgggggcg gcggcgccat cttgtgcccg gggccggtgg
1381 ggagggccggg gagggggccc cgggggggcgc agggggactac gggaacggccc tggagtctga
1441 ggaactggag cctgagggagc tgctgctgga gcccgagccg gagcccgagc ccgaagagga
1501 gccgccccgg ccccgcgccc ccccgggagc tccgggccct gggcctggtt cgggagcccc
1561 cggcagccaa gaggaggagg aggagccggg actggtcgag ggtgacccgg gggacggcgc
1621 cattgaggac ccggtgagga aggagggcga gcgagcaggc cggcggctgg cgcgtcactg
1681 gaggcccaga gctcgggcga gcggtggcag gcgggggggtg gggttgggcg gggaataacg
1741 tggctgggggc gggtcgggccc ggggatgggt cagcgatcac tacaaggggc ccgactggct
1801 tgattcgggtc gtcacgggtg cctagtgttg ttctagagag ggtagctttt ctttatcac
1861 gaccctcgca tgggggcgagg gaaatggccg agcatggctg aggcgcgctc tggccgagag
1921 cagggcacag ccctgcgtt ggttcctctt aagctgtcct ccatacccctc cccacttata
1981 ttaggagctg gaagctatca aagctcgagt cagggagatg gaggaagaag ctgagaagct
2041 aaaggagcta cagaacgagg tagagaagca gatgaatatg agtccacctc caggcaatgc
2101 tgagtaactg gcggttgcac gcggagcccg ggttctcggg ttggaagggt tgtggggagg
2161 atgggggaatg tgggggttaga tactcggcac cctggagctg cttgtctgag ctattatgac
2221 tgtgccgcgg tcatagtccg ttgtgtgttc ctctgaccct tgtgaggcag aactgatatt
2281 ttggtggtgg tagccttgtg cctcccttttg tcctgttata attgtgttgc tctttattct
2341 tagtctacgt ctatctttct ttggtagagg ttgcgtgctc gcatttgacc ttcaaatcta
```

Fig. 4B

```
2401 atagtttttc ctccaattgg agacgcttta ggattctaag agaaagcaag ctggaagggg
2461 ttcccccttt aaattctaga aatgtggagt ctcagcccac ttaatttttgc tcactcttaa
2521 aagcatttca accaaagcca ttcattaggg atttgatttg gagggcagga gggattccta
2581 tactgttttta agtgtgtatt aattctttca atttatcgaa ttatttagtg agtaacctgc
2641 tatgcactag gcactattct cggcttgtgg gtacagcagg gaacagcaca gaccaaaatc
2701 tttgccttca ctgagcttat gggatagtgc tggtggtgga agtgcaacat attggtcaag
2761 tagaaaacaa gtgtgtggtt tttgtaaaaa attatttttt cctgatagct ggcccggtga
2821 tcatgtccat tgaggagaag atggaggctg atgcccgttc catctatgtt ggcaatgtga
2881 cgtactgggg ctctgactgg ggttgggggc aagttcttct tttgggggaat tatttaatag
2941 tcctgaaaga acatctccgg gatagatgtg gttttgggtg tggagggagt gtgggaagga
3001 ggttaaaggt aatggaatga tcagtaatca gcaaaggctc tgggtttgga aggaaaagag
3061 attaattcct caaattacca gatttcatgt gctttggtgt atgatggccc agaccaaagg
3121 ctcgggaggg ttcttttgag acaggaattt gcctggtgcc tgtgaaattt ttctccctctc
3181 atcaggtgga ctatggtgca acagcagaag agctggaagc tcactttcat ggctgtggtt
3241 cagtcaaccg tgttaccata ctgtgtgaca aatttagtgg ccatcccaaa ggtaaagtaa
3301 aggggagtaa gttgagataa tttaaattac agtgtacaaa tagataaatt atgttttata
3361 ttgagcagta agttatttgg tgttaacaca ggtgatctgt gtcatttaag atcatggcat
3421 taatgttgat atatcaggag ttgcacctaa atgtcttcag aggccagata acaaaaatga
3481 aggctagatg tgggtgggat tacgaactag aaggggaggg gcagcttcta cttggcctat
3541 tatggcatat ggaaattcag gccctgtgtg tcttatttct acaaatttca aagagtagct
```

FIG - 4C

```
3601 ggaaatttta aaatttaaat gatttcgaat gattgaaatt ttccatttag aagaattttg
3661 acaaataaaa aatataactg cattgtagcc caaaacgaag catgcctgca ggttgaattt
3721 gacctgtgag gtatttgtaa cctcagagag atacaatgac aattctttcc aggtttgcgt
3781 atatagagtt ctcagacaaa gagtcagtga ggacttcctt ggccttagat gagtccctat
3841 ttagaggaag gcaaatcaag gtaagcctat gtccattgct gttctagttg tgtataaact
3901 ctccaggttg cctttaaggc tatcatttgt tcatctctga ctcaggtgat cccaaaacga
3961 accaacagac caggcatcag cacaacagac cggggttttc cacgagcccg ctaccgcgcc
4021 cggaccacca actacaacag ctcccgctct cgattctaca gtggttttaa cagcaggccc
4081 cggggtcgcg tctacaggtc aggatagatg ggctgctcct ctttcccccg cctcccgtga
4141 gccccgtatg cttcctcctc tctggtctga ggaacctccc tccccccacc cctccccgtg
4201 gtcttcagga actttgtctc ctgcctgtgc aggttgagga aggtagtttgc aggccaggcc
4261 agaaggcagc ctcatcatct tttctgcagt agaaaattggt gataagggct gcatccctcc
4321 cttggttcaa agaggcttcc acccccagcc ttttttttct tgggagttgg tggcatttga
4381 aggtgtttgc ggacaaaact gggagggaaca gggcctccag gaagttgaaa gcactgcttg
4441 gacatttgtt actttttcg gagttaggga gggattgaag actgaacctc ccttggaaga
4501 ataccagagg ctagctagtt gatcctccca acagccttgt gggaggattt tgagatactt
4561 attctttatt tgagccagtc ttgcaaggtt aacttctcac tgggcctagt gtggtnccca
4621 ggtttttgcc ttgcttcact tctgtctcta catttaaata gacgggttag gcatataaac
4681 cttggctttt cataagctct acctgcctat ccccaggagt tagggaggat ctatttgtga
4741 aggccctagg gtttaaaaac tgtggaggac tgaaaaactg gataaaaagg gggtccttct
```


Fig - 4D

```
4801 ccttgcccct gtctctcact cagatgcgct tctttttcgc cactgtttgg caaagttttc
4861 tgttaagccc ccctcccccc gccccagttc tcccaggtgc gttactattt ctgggatcat
4921 ggggtcggtt ttaggacact tgaacacttc ttttcccccc ttcccttcac agtaactggg
4981 gcaggggcct acggggaggg gcttgtactg aactatctag tgatcacgtt aacacctaac
5041 tctccttctt tcttccaggg gccgggctag agcgacatca tggtattccc cttactaaaa
5101 aaagtgtgta ttaggaggag agagaggaaa aaaagaggaa agaaggaaaa aaaaaagaat
5161 taaaaaaaaa aaaaagaaaa acagaagatg accttgatgg aaaaaaaata ttttttaaaa
5221 aaaagatata ctgtggaagg ggggagaatc ccataactaa ctgctgagga gggacctgct
5281 ttggggagta ggggaaggcc cagggagtgg ggcagggggc tgcttattca ctctggggat
5341 tcgccatgga cacgtctcaa ctgcgcaagc tgcttgccca tgtttccctg ccccccttcac
5401 ccccttgggc ctgctcaagg gtaggtgggc gtgggtggta ggagggtttt ttttacccag
5461 ggctctggaa ggacaccaaa ctgttctgct tgttaccttc cctcccgtct tctcctcgcc
5521 tttcacagtc ccctcctgcc tgctcctgtc cagccaggtc taccacccac cccacccctc
5581 tttctccggc tccctgcccc tccagattgc ctggtgatct attttgtttc cttttgtgtt
5641 tctttttctg ttttgagtgt ctttctttgc aggtttctgt agccggaaga tctccgttcc
5701 gctcccagcg gctccagtgt aaattccccc tccccctggg gaaatgcact accttgtttt
5761 gggggggttta ggggtgtttt tgtttttcag ttgttttgtt ttttgtttt ttttttttcc
5821 tttgcctttt ttccctttta tttggaggga atgggaggaa gtgggaacag ggaggtggga
5881 ggtggatttt gtttattttt ttagctcatt tccaggggtg ggaatttttt tttaatatgt
5941 gtcatgaata aagttgtttt tgaaaataaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
6001 aa
```

FIG. 4E

SEQUENCE LISTING

<110> McGILL UNIVERSITY
ROULEAU, Guy A.
BRAIS, Bernard

<120> SHORT GCG EXPANSIONS IN THE PAG II GENE FOR OCULOPHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC THEREOF

<130> 1770-199PCT FC/ld

<150> CA 2,218,199
<151> 1997-12-09

<160> 6

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 57
<212> DNA
<213> Artificial Sequence

<400> 1
atggcggcgg cggcggcggc ggcagcagca atggcggcgg cggcggcggc ggcggca 57

<210> 2
<211> 35
<212> DNA
<213> Artificial Sequence

<400> 2
maaaaaaaaa agaaggrgsm aaaaaaaaaa agaag 35

<210> 3
<211> 6002
<212> DNA
<213> Artificial Sequence

<400> 3
aatgaaggtg gacacccaaa tagccccaat acaaatgcct gttcaatcaa ccaaacatct 60
aagcagcaca tctatgtggt agcatattgc caggccgtga gactgcgaat ataaatagga 120
accgcccctc atctgcaggc gctcacaacc tagttagcaa acagtaaaac aattaagcgc 180
gccgtggaca taggcccact tgtcctggga aatgagggga agctggggtt tgcagtggtt 240
tgattgaagg gggactacat gttagaggca cagactgggt gcaggtacac ccaaaggaac 300
gagaagagtg gaaggaaaca acatccacaa agtaaccaca tgctggcgta tcgaaggccg 360
tgatttacgg ttttgagact ttacctcgcc agcaaagggg ggccagtctg ttagcggtgc 420
agattggagg ggtgacattg gaagctgtcc aggaaaaaga aaatggaact ggggagcaga 480
aggcctacgc aagagggcgg gacagacagg acttgtgact agtagctctg gactgaggaa 540
tcctccctgc tttctggtgc gggagagcta gtggatgatg gtgccaataa cctggatggg 600
gaaagtaagc tccctcctgg aatgcttcat tcacaacctc cattttcagc aacatcccat 660
ctactggtgc ttcctggtcg agatacaagt ttcctgaaac tgctgctctg ttttgggcct 720
cacccggcca acagctcact agctggcaag cagtagtatc aagatggcgg ccccctagga 780
ctggctagtc atgtgacctc gggtttccca agtttgaagc ccggcagtcc tttcgggggc 840
aaggttcacc tgtcacgaaa cgagtgtcac cccttcgact ctcgcaagcc aatcggcatc 900
tgagactggg ccactgcggt gaggcgatcg gaagattggt cctttccagt cgcctagcta 960

gggccaatca cggagcgtcc catacttcgc gggcccgccc gtaggccggg gagaagcagg 1020
aatatcgtca cagcgtggcg gtattattac ctaaggactc gataggaggt gggacgcgtg 1080
ttgattgaca ggcagatttc cctaccggga tttgagaatt tggcgcagtg cccgccttag 1140
aggtgcgctt atttgattgc caagtaatat tccccaatgg agtactagct catggtgacg 1200
ggcaggcagc ttgagctaat gagtcctccg tggccggcgc agctctccac atgccgggcg 1260
gcgggcccca gtctgagcgg cgatggcggc ggcggcggcg gcggcagcag cagcgggggc 1320
tgcgggcggt cggggctccg ggccggggcg gcggcgccat cttgtgcccg gggccggtgg 1380
ggaggccggg gaggggggcc cgggggggcgc aggggactac gggaacggcc tggagtctga 1440
ggaactggag cctgaggagc tgctgctgga gcccgagccg gagcccgagc ccgaagagga 1500
gccgccccgg ccccgcgccc ccccgggagc tccgggccct gggcctggtt cgggagcccc 1560
cggcagccaa gaggaggagg aggagccggg actggtcgag ggtgacccgg gggacggcgc 1620
cattgaggac ccggtgagga aggagggcga gcgagcaggc cggcggctgg cgcgtcactg 1680
gaggcccaga gctcgggcga gcggtggcag gcgggggtg gggttgggcg gggaataacg 1740
tggctggggc gggtcgggcc ggggatgggt cagcgatcac tacaaggggc ccgactggct 1800
tgattcgggc gtcacgggtg cctagtgttg ttctagagag ggtagctttt cttttatcac 1860
gaccctcgca tggggcgagg gaaatggccg agcatggctg aggcgcgctc tggccgagag 1920
cagggcacag cccctgcgtt ggttcctctt aagctgtcct ccataccctc cccacttata 1980
ttaggagctg gaagctatca aagctcgagt cagggagatg gaggaagaag ctgagaagct 2040
aaaggagcta cagaacgagg tagagaagca gatgaatatg agtccacctc caggcaatgc 2100
tgagtaactg gcggttgcac gcggagcccg ggttctcggg ttggaagggt tgtggggagg 2160
atggggaatg tggggttaga tactcggcac cctggagctg cttgtctgag ctattatgac 2220
tgtgccgcgg tcatagtccg ttgtgtgttc ctctgacctt tgtgaggcag aactgatatt 2280
ttggtggtgg tagccttgtg cctccctttg tcctgttata attgtgttgc tctttattct 2340
tagtctacgt ctatctttct ttggtagagg ttgcgtgctc gcatttgacc ttcaaatcta 2400
atagtttttc ctccaattgg agacgcttta ggattctaag agaaagcaag ctggaagggg 2460
tttccccttt aaattctaga aatgtggagt ctcagcccac ttaattttgc tcactcttaa 2520
aagcatttca accaaagcca ttcattaggg atttgatttg gagggcagga gggattccta 2580
tactgtttta agtgtgtatt aattctttca atttatcgaa ttatttagtg agtaacctgc 2640
tatgcactag gcactattct cggcttgtgg gtacagcagg gaacagcaca gaccaaaatc 2700
tttgccttca ctgagcttat gggatagtgc tggtggtgga agtgcaacat attggtcaag 2760
tagaaaacaa gtgtgtggtt tttgtaaaaa attatttttt cctgatagct ggcccggtga 2820
tcatgtccat tgaggagaag atggaggctg atgcccgttc catctatgtt ggcaatgtga 2880
cgtactgggg ctctgactgg ggttgggggc aagttcttct tttggggaat tatttaatag 2940
tcctgaaaga acatctccgg gatagatgtg gttttgggtg tggagggagt gtgggaagga 3000
ggttaaaggt aatggaatga tcagtaatca gcaaaggctc tgggtttgga aggaaaagag 3060
attaattcct caaattacca gatttcatgt gctttggtgt atgatggccc agaccaaagg 3120
ctcgggaggg ttcttttgag acaggaattt gcctggtgcc tgtgaaattt ttctcctctc 3180
atcaggtgga ctatggtgca acagcagaag agctggaagc tcactttcat ggctgtggtt 3240
cagtcaaccg tgttaccata ctgtgtgaca aatttagtgg ccatcccaaa ggtaaagtaa 3300
aggggagtaa gttgagataa tttaaattac agtgtacaaa tagataaatt atgttttata 3360
ttgagcagta agttatttgg tgttaacaca ggtgatctgt gtcatttaag atcatggcat 3420
taatgttgat atatcaggag ttgcacctaa atgtcttcag aggccagata acaaaaatga 3480
aggctagatg tgggtgggat tacgaactag aaggggaggg gcagcttcta cttggcctat 3540
tatggcatat ggaaattcag gccctgtgtg tcttattttt acaaatttca aagagtagct 3600
ggaaatttta aaatttaaat gatttcgaat gattgaaatt ttccatttag aagaattttg 3660
acaaataaaa aatataactg cattgtagcc caaaacgaag catgcctgca ggttgaattt 3720
gacctgtgag gtatttgtaa cctcagagag atacaatgac aattcttttc aggtttgcgt 3780
atatagagtt ctcagacaaa gagtcagtga ggacttcctt ggccttagat gagtccctat 3840
ttagaggaag gcaaatcaag gtaagcctat gtccattgct gttctagttg tgtataaact 3900
ctccaggttg cctttaaggc tatcatttgt tcatctctga ctcaggtgat cccaaaacga 3960
accaacagac caggcatcag cacaacagac cggggttttc cacgagcccg ctaccgcgcc 4020
cggaccacca actacaacag ctcccgctct cgattctaca gtggttttaa cagcaggccc 4080
cggggtcgcg tctacaggtc aggatagatg ggctgctcct ctttcccccg cctcccgtga 4140
gccccgtatg cttcctcctc tctggtctga ggaacctccc tccccccacc cctccccgtg 4200
gtcttcagga actttgtctc ctgcctgtgc aggttgagga aggtagttgc aggccaggcc 4260
agaaggcagc ctcatcatct tttctgcagt agaaattggt gataagggct gcatccctcc 4320
cttggttcaa agaggcttcc acccccagcc ttttttttct tgggagttgg tggcatttga 4380

aggtgtttgc ggacaaaact gggaggaaca gggcctccag gaagttgaaa gcactgcttg 4440
gacatttgtt actttttcg gagttaggga gggattgaag actgaacctc ccttggaaga 4500
ataccagagg ctagctagtt gatcctccca acagccttgt gggaggattt tgagatactt 4560
attctttatt tgagccagtc ttgcaaggtt aacttctcac tgggcctagt gtggtnccca 4620
ggtttttgcc ttgcttcact tctgtctcta catttaaata gacgggttag gcatataaac 4680
cttggctttt cataagctct acctgcctat ccccaggagt tagggaggat ctatttgtga 4740
aggccctagg gtttaaaaac tgtggaggac tgaaaaactg gataaaaagg gggtcctttt 4800
ccttgcccct gtctctcact cagatgcgct tctttttcgc cactgtttgg caaagttttc 4860
tgttaagccc ccctcccccc gccccagttc tcccaggtgc gttactattt ctgggatcat 4920
ggggtcggtt ttaggacact tgaacacttc ttttcccccc ttcccttcac agtaactggg 4980
gcaggggcct acggggaggg gcttgtactg aactatctag tgatcacgtt aacacctaac 5040
tctccttctt tcttccaggg gccgggctag agcgacatca tggtattccc cttactaaaa 5100
aaagtgtgta ttaggaggag agagaggaaa aaaagaggaa agaaggaaaa aaaaaagaat 5160
taaaaaaaaa aaaaagaaaa acagaagatg accttgatgg aaaaaaaata ttttttaaaa 5220
aaaagatata ctgtggaagg ggggagaatc ccataactaa ctgctgagga gggacctgct 5280
ttggggagta ggggaaggcc cagggagtgg ggcagggggc tgcttattca ctctggggat 5340
tcgccatgga cacgtctcaa ctgcgcaagc tgcttgccca tgtttccctg ccccccttcac 5400
ccccttgggc ctgctcaagg gtaggtgggc gtgggtggta ggagggtttt ttttacccag 5460
ggctctggaa ggacaccaaa ctgttctgct tgttaccttc cctcccgtct tctcctcgcc 5520
tttcacagtc ccctcctgcc tgctcctgtc cagccaggtc taccacccac cccacccctc 5580
tttctccggc tccctgcccc tccagattgc ctggtgatct attttgtttc cttttgtgtt 5640
tctttttctg ttttgagtgt ctttctttgc aggtttctgt agccggaaga tctccgttcc 5700
gctcccagcg gctccagtgt aaattcccct tccccctggg gaaatgcact accttgtttt 5760
ggggggttta ggggtgtttt tgtttttcag ttgttttgtt tttttgtttt ttttttttcc 5820
tttgcctttt ttccctttta tttggaggga atgggaggaa gtgggaacag ggaggtggga 5880
ggtggatttt gtttattttt ttagctcatt tccaggggtg ggaatttttt tttaatatgt 5940
gtcatgaata aagttgtttt tgaaaataaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 6000
aa 6002

<210> 4
<211> 19
<212> DNA
<213> Artificial Sequence

<400> 4
cgcagtgccc cgccttaga 19

<210> 5
<211> 24
<212> DNA
<213> Artificial Sequence

<400> 5
acaagatggc gccgccgccc cggc 24

<210> 6
<211> 23
<212> DNA
<213> Artificial Sequence

<400> 6
cgcagtgccc cgccttagag gtg 23

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/CA 98/01133

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68 A01K67/027 C12N15/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q A01K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NEMETH A ET AL: "Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II." NUCLEIC ACIDS RESEARCH, (1995 OCT 25) 23 (20) 4034-41, XP002097093 see the whole document ---	1-5
Y	DATABASE EMBL Entry HS12206 Accession Nr. U12206, 4 March 1995 XP002097097 see abstract ---	1-5
Y	DATABASE EMBL Entry AA618589 Accession Nr. AA618589, 10 October 1997 XP002097098 see abstract ---	1-5
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 March 1999	01/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Reuter, U

4

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter .onal Application No
PCT/CA 98/01133

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	TEH B T ET AL: "Oculopharyngeal muscular dystrophy (OPMD)--report and genetic studies of an Australian kindred." CLINICAL GENETICS, (1997 JAN) 51 (1) 52-5, XP002097094 see the whole document	1-10
A	BRAIS B ET AL: "Using the full power of linkage analysis in 11 French Canadian families t fine map the oculopharyngeal muscular dystrophy gene." NEUROMUSCULAR DISORDERS, (1997 OCT) 7 SUPPL 1 S70-4, XP002097095 see page 70 see page 74	1-10
A	RIGGINS G J ET AL: "HUMAN GENES CONTAINING POLYMORPHIC TRINUCLEOTIDE REPEATS" NATURE GENETICS, vol. 2, no. 3, November 1992, pages 186-191, XP000196593 see the whole document	1-10
P,X	BRAIS B ET AL: "Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy" NATURE GENETICS, (1998 FEB) 18 (2) 164-7 , XP002097096 see the whole document	1-10
P,X	WO 98 31800 A (KRISSANSEN GEOFFERY W ;AUCKLAND UNISERVICES LTD (NZ); HUMAN GENOME) 23 July 1998 see page 67 - page 68; claims 1-3	1-5

4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA 98/ 01133

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 10 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 11
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Claim 11 is obscure. The mentioned hGT1 gene is not mentioned in claim 1, and claim 11 lacks technical details.

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

...ormation on patent family members

Interr ...al Application No
PCT/CA 98/01133

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9831800 A	23-07-1998	AU 5923398 A	07-08-1998
		AU 6030398 A	07-08-1998
		WO 9831799 A	23-07-1998
		AU 5927398 A	07-08-1998
		AU 6241698 A	07-08-1998
		WO 9831818 A	23-07-1998
		WO 9831801 A	23-07-1998
		WO 9831806 A	23-07-1998